

10 Reasons to Invest in On-Prem AI

Flexibility, performance, security and more

From enhanced data storage and unique customizations to increased operational efficiency and infrastructure scalability, there are more reasons than ever to invest in on-prem AI. Discover how Cisco, NVIDIA and ConRes can accelerate your AI journey.

1 Superior Performance

Your AI applications run faster and more efficiently, giving you a competitive edge.

Cutting-edge processor acceleration like NVIDIA's GPUs and efficient software frameworks maximize system performance. Plus, repurpose this hardware for other tasks when AI isn't in demand.



2 Tailored Customization

Greater efficiency and better results with solutions tailored to your needs.

Optimize AI deployment by customizing your on-prem infrastructure to suit your specific organizational workloads, enhancing performance and efficiency.

3 Seamless Scalability

Easily adapt to increasing demands without disrupting operations.

As your AI workloads grow, your infrastructure should scale effortlessly. On-prem solutions based on Cisco Validated Designs can reduce unplanned downtime by nearly 90%¹, allowing smooth transitions from a single workstation to an AI cluster.

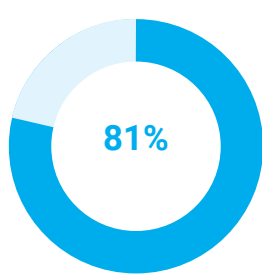
4 Enhanced Data Privacy

Deploy your AI projects while keeping sensitive data secure and under your control.

With 82% of data breaches involving cloud-stored data², robust on-prem security measures can better safeguard your information from unauthorized access and cyber threats, giving you maximum control.

5 Operational Effectiveness

Efficient and effective AI deployments (finally).



81% of IT leaders say that data silos are hindering robust AI adoption.³ By storing your data and deploying AI infrastructure in the same location, you can reduce those silos. Cisco's AIOps, combined with NVIDIA and OpenShift, streamlines the management and control of IT functions across AI applications.

6 Reduced Latency

Improved speed and user experience, plus optimal performance.

Latency levels above even just a few seconds can significantly hinder the adoption of any AI-based application.⁴ Avoid performance lags with on-prem infrastructure, where AI operations run smoothly with minimal data latency.

7 Efficient Data Storage

Effective and predictable data storage costs.

Managing cloud spend is a significant challenge facing organizations.⁵ On-prem infrastructure lets you scale your storage needs upfront, helping avoid escalating subscription costs from cloud-based AI storage.

8 Long-Term Cost Effectiveness

An optimized budget that enables AI adoption while saving money in the long run.

While on-prem infrastructure might seem costly upfront, it proves more economical over time. Over 60% of SMBs spend more than \$600k on cloud computing.⁵ On-prem solutions are the clear economic choice for companies planning to use AI in a day-to-day fashion, or if they plan on using large amounts of in-house data to fuel their AI solutions. In short, on-prem AI infrastructure, designed to scale if needed is a smart choice.



9 Stronger Organizational Security

An enhanced security posture that reduces risks without hindering innovation.

On-prem AI allows organizations the flexibility to control their data and to customize the security policies protecting it. This also can also assure regulatory and compliance requirements are met.



10 Streamlined Software Integration

Seamless integration and maximized AI capabilities.

Combine cloud-based software with powerful on-prem infrastructure to fully realize the potential of hybrid solutions. This integration facilitates building and testing models, APIs and data repositories.

Accelerate your AI journey with ConRes, Cisco and NVIDIA. From discovery through deployment, our expertise in on-prem AI infrastructure helps you streamline operations, enhance security and improve cost savings. To get started, contact us at solutions@conres.com.

1 <https://www.cisco.com/c/dam/en/us/solutions/collateral/enterprise/data-center-designs-cloud-computing/reduce-risk-customers.pdf>
 2 <https://www.ibm.com/downloads/cas/E3G5JMBP>
 3 <https://www.mulesoft.com/lp/reports/connectivity-benchmark>
 4 <https://www.forbes.com/sites/sylvainduranton/2024/02/05/the-bermuda-triangle-of-generative-ai-cost-latency-and-relevance>
 5 <https://info.flexera.com/CM-REPORT-State-of-the-Cloud-2024-Thanks>